

# Returning Home and Worsening the War: The Causal Effect of Refugee Return on Civil Conflict Intensity

Kara Ross Camarena

This Draft: January 2016

## **Abstract**

As Syrian refugees continue to flee from their ongoing civil war, the countries neighboring Syria warn of the refugees' destabilizing effects in their countries of asylum. Political scientists have focused on rebels and militants hidden amongst the refugees as the drivers of the spreading violence. However, large population movements are associated with increases in violence, even in the absence of civil wars and rebel groups. In fact, a collection of recent work in sociology has re-examined the connection between violent crime in American cities during large waves of immigration, and concluded that it is not the individual migrants who are responsible for the increase. Rather scholars point to the structural and institutional changes precipitated by the large population increases as the underlying cause. Similarly, large refugee movements because of their size, can overextend institutions in the country of asylum, leading to violence. It is difficult though to know whether increases in violence that are caused by large refugee influxes are because of the militants among the refugees or because of the population shift more generally. I propose a novel identification strategy for disentangling these sources of violence. Focusing on the situation in which refugees return during an ongoing civil war, I use natural disasters as an instrument. I estimate that violence increases 3%-8% for every 100,000 returnees. The measurable increase in violence in the context of civil war gives credence to concerns over large population movements destabilizing their destination, not because those arriving are strategic violent agents, but because the existing structures that maintain order fail.

## Introduction

In 2014 US policy advisors and the Prime Minister of Lebanon himself publicly worried that the influx of Syrian refugees, along with other spillover effects from the Syrian conflict, would destabilize Lebanon (Mikati 2014; Yacoubian 2014). The message was that Syrian refugees, coming in such large numbers, had begun to overwhelm Lebanon. The sheer numbers support such a claim, the Syrian refugee population in Lebanon is more than a third the population of Lebanon before the current Syrian crisis began.<sup>1</sup> The Syrian refugee influx continues, and taxes Lebanon's ability to provide basic needs to people in great need fleeing the Syrian civil war. Perhaps more problematically, the refugees have become an issue in domestic politics, raising sectarian tensions in Lebanon (ICG 2013; Yacoubian 2014) .

The concern that a mass influx of refugees will destabilize a country of asylum is not a new one. Massive shifts or influxes in population, along with people in desperate situations, often are associated with increases in violence. Refugees destabilizing the Balkans, and in particular Macedonia, was a topic of concern following the break up of Yugoslavia (ICG 1999). Rwandan refugees fleeing the 1994 genocide spurred violence that led to a civil war in the Democratic Republic of the Congo. Liberian refugees were a critical factor in the destabilization and civil war in Sierra Leone (Saleyan 2011). Even earlier in history, policy advisors worried desperate victims of war would prompt social unrest. The origins of the United States' foreign food aid program can be traced to preventing violence and migration and addressing the humanitarian crisis in occupied Belgium during World War I (Lucey 1914).

There need not be a war at all for population movements to be associated with increases in violence. A substantial amount of scholarly work in sociology examines how South-North

---

<sup>1</sup>Estimates suggests that nearly 1.5 million Syrian refugees are living in Lebanon (UNHCR 2015). Lebanon's population in 2010 before the Syrian conflict was 4.3 million (World Bank 2010).

migration and European Immigration lead to increased violent crime in American cities throughout the first part of the 20th century (For bookends see [Shaw and McKay \(1942\)](#) and [Sampson and Laub \(2005\)](#) ). This work highlights that even in the context of strong institutions to maintain order, large population movements are associated with violent events. Nevertheless, much of the literature in political science on refugee influxes and subsequent violence or civil war has hinged on the presence of rebels or at least militantly leaning refugees among the refugees. In the wake of the Rwandan genocide and the complex set of fighters among refugees in the DRC, the issue of militants hiding among refugees was particularly salient ([Lischer 2006](#)). Scholars of international relations have explored issues of refugees, transnational rebels, and the regional clustering of civil wars. These works to some extent acknowledge failures in basic institutions due to the demands of the large increases in population. Further, these institutional failures may help explain how refugee influxes contribute to civil wars in their destination. Still though, these scholars have largely focused on identifying the agents of violence, rebels within the refugee movements ([Salehyan 2008](#); [Salehyan and Gleditsch 2006](#)). As a result in the associated empirical work, estimates conflate these two sources of violence—violence associated with large populations movements and violence instigated by rebels moving with refugees across borders.

I am able to get leverage on the question to what extent do large population movements, apart from rebels and militants, account for subsequent violence by thinking about a slightly different context. One logical extension of the notion that refugees spread conflict to their destination is that when they return to their country of origin they should increase violence in the conflict they had once fled. It could be the case that refugees become more militarized while away or because when they return home they represent the same kind of demand shock to their home country that they did to the asylum country when they first arrived.

Such a relationship may be difficult to unpack empirically though. This is because most

refugees return home because they believe it is safe to do so. They only return when there is already a downward trend in violence. Thus the extent to which refugee return causes an increase in violence is likely overshadowed in the data by the expected downward trend in violence.

To examine this relationship, I make use of the exogenous variation in natural disasters in asylum countries and instrument for refugee return. I find that the relationship between refugee return and increased violence in civil conflicts is negative but insignificant in an ordinary least squares regression (OLS). By contrast, the two stage least squares (2SLS) regression reveals a positive, significant relationship between refugee return and increased violence. In other words, when I account examine spontaneous return, rather than return in response to an ending war, I see an increase in violence. Furthermore, this positive relationship persists even when controlling for characteristics traditionally thought to explain variation in civil war including ethnic fractionalization, GDP per capita, and rough terrain.

The paper proceeds as follows. First, I describe the causal relationship of interest, the challenges to identification, and my empirical strategy. Second, I outline the particulars of the data used in the analysis. Third I present the results of the primary analysis and highlight the results of some robustness checks. Last I explore the results of the analysis in the context of existing theory and suggest some avenues for additional research.

## **The Causal Relationship: Refugee Flows and Violence**

In this section, I briefly outline the theoretical ideas that underpin some of the empirical studies on how refugee flows may be associated with subsequent civil war in countries of asylum. Then, I explore the difficulty with interpreting the results of the empirical research causally. I then explain how my approach addresses these and other challenges. Finally, I explicitly lay out my identification strategy.

Collectively, the scholarship on transnational civil war and conflict contagion has argued that armed conflict can be spread through refugee flows. There are essentially three mechanisms explored in the literature. First, refugees may be the reason for international disputes either because one country pursues its refugees across an asylum country's border or because a potential asylum country uses its military to prevent refugees from entering (Gleditsch et al. 2008). Second, refugees often present additional economic, administrative, and security burdens on an asylum country's governments, and this leads to destabilization (Salehyan and Gleditsch 2006). Last, militarized refugees may connect with like-minded asylum country nationals, broadening the scope of a rebel movement to include an asylum country (Salehyan and Gleditsch 2006). In this project, I set aside the first explanation, and distinguish the second from the third. In other words, I disaggregate what violence is attributable to the population shock and the violence attributable to the rebels themselves.

While the empirical literature clearly establishes an association between refugee flows and subsequent violence, these empirical analyses do not explicitly take into account the possibilities of alternative causal relationships. One possibility is that reasons that refugees move to a particular asylum country are simultaneously the reasons the asylum country is more prone to conflict. More precisely, the selection of the asylum country and the civil war have a common cause. For example, Rwandans may flee to Burundi not just because of its proximity, but because Burundi has similar ethnic groups, language, history and way of life. It is not altogether clear whether it is these similar characteristics or the presence of the refugees that cause the armed conflict.

A second problem is that it is possible that the conditions that cause a civil war are present in two countries prior to conflict or refugee flows, that is, that civil war in the origin country and asylum country are caused by the same thing. In this scenario, one country's civil war begins first and refugees flee to the second country. Then, the second country's civil war

begins and we infer, incorrectly, a causal relationship, a spreading of the civil war. In fact, the civil war in the second country, as I have indicated above, may have been caused because of similar reasons as the first country, not as a result of spillover or contagion.

Finally, it is sometimes the case in the data that countries are simultaneously exchanging refugees, that is refugees from one country are living in a second country, while at the same time, refugees from the second country are living in the first country. This scenario highlights the potential endogenous cycle in the data. Specifically, the explanation that refugees bring violence to their country of asylum is potentially happening in two directions. Refugees spread violence from one country to a second, while refugees from the second country spread violence to the first.

## **Focusing on Returnees**

I am able to address the identification problems of earlier work, by looking at a slightly different situation with an explicit strategy to get around the cycle problem. First, rather than looking at conflict onset in an asylum country following refugee flows, I study increases in violence when refugees return home during an already ongoing civil war. By changing the context and looking at changes in violence with returnees, I focus on a single country and eliminate the common cause concerns.

However, using returnees is not without problems that need to be addressed. Namely, it is generally believed that refugee return occurs with diminished violence, which suggests the opposite causal relationship between violence and returnees. More particularly there two related reasons to expect violence should go down with refugee return. The first is there is likely already a downward trend in violence when returnees begin arriving home because they wait until they believe it is safe before they return to their country of origin. In general refugee repatriation occurs after hostilities have ended, and violence has decreased.

Less violence, makes it more likely that the refugees will be able to return safely, settle and establish a livelihood for themselves.

The second reason to expect a negative relationship attributes a bit more sophistication to the refugees. This explanation posits that refugees have good information about the situation in their home country. Because of this, they can anticipate when violence will decrease sufficiently to return home and therefore, decreased violence will follow or occur with their return home. By either of these two logics there should be negative relationship between violence and refugee return, not because returnees decrease the violence in their country of origin, but rather because the returnees are responding to or anticipating decreased violence by coming home. This means that a naive approach to estimating the impact of refugee return on violence is likely to be negative because refugees are responding to resolution of the violence and returning home.

## **Random Variation in Natural Disasters**

I harness the random variation in natural disasters in order identify increases in violence caused by refugee return. This empirical approach accomplishes three things: first, because the natural disaster is random, I need not be concerned about refugees causing violence only to cause more refugees and more violence in the data. I know the returnees are moving because of a random event, not violence that can be endogenous to other violence and refugee flows. Second, these refugees are returning home because of the natural disaster not because of expectations or the realization of the resolution of the civil war. Therefore, I no longer have two causal stories that go in opposite directions. Third, since neither the rebel group nor the refugees could have anticipated the natural disaster, the instrument identifies the violence associated with the population movements rather than some organized return of the rebels in exile.<sup>2</sup>

---

<sup>2</sup>Some have raised the concern that rebels could strategically wait for a natural disaster as cover for their return. I have found no account in a set of cases in Africa that I am familiar with where anyone suggests this

To understand why natural disasters make a good instrument, I need to explain two things, (1) why natural disasters in the country of asylum drive refugees back to their country of origin, (2) why natural disasters in the country of asylum is a valid instrument.

First, a simple decision making model of the refugee decision to move explains why a natural disaster in the asylum country pushes refugees back to their country of origin. Prior to the natural disaster, the refugee prefers to live in the country of asylum as compared to the country of origin. However, following a significant natural disaster, the quality of life in the country of asylum may decrease, making it a relatively less good place to live than before the natural disaster. If the decrease in the quality of life for the refugee is sufficiently large, the refugee may choose to incur the cost of moving to live in a relatively better place.

Now the refugee must choose between moving home and moving to some third country. There are three reasons that returning home might be optimal. The first two are related to the cost of moving. In particular first, refugees often live fairly close to the border with their country of origin, meaning their country of origin is the closest country to which to move. Second, because of proximity and connections back home information about moving to and the conditions of the country of origin is often easier to obtain than a country where the refugee has few connections and has never visited. The third reason return may be optimal is the benefits of returning to the country of origin. It is often the case for a host of reasons that people derive utility from returning home to a place they once belonged, making this option more appealing than going to some other country where they will be a stranger.

Second, for natural disasters to be a valid instrument, it must satisfy the exclusion restriction.

---

has occurred. Furthermore, it seems heroic to think that rebels are using natural disasters as cover for return, strategically, when they cannot know whether a large enough natural disaster will occur any time in the near future. However, to be clear, I argue that there is no reason to think that rebels will systematically wait for sufficiently large natural disasters and then also effectively anticipate large populations of returnees, so that they can launch an attack. Rather to the extent that a rebel group has leveraged a natural disaster in their fight, this is spurious and not systematic. Finally, the empirical results on natural disasters discussed at the end of this article reinforce the credibility of my argument.

Namely, there should be no reason to think a natural disaster in the asylum country causes violence in the country of origin except because of refugee return. Since natural disasters are not bound by country borders I ensure this by including controlling for natural disasters in the country of origin. Furthermore, since levels of rainfall has been linked to violence and droughts often encompass entire regions, I explicitly exclude droughts from my possible natural disasters.

Before specifying the empirical strategy in detail in the section I summarize the logic of my empirical strategy. First, since the returnees of interest are individuals coming back to their country of origin in the midst of an ongoing civil war, I avert the common cause identification problems of studying the impact of refugee flows on violence in their country of asylum. Second since I make use of random variation in natural disasters, I eliminate concerns about refugee-violence-refugee cycle. Also harnessing the random variation in natural disasters I am able to identify spontaneous refugee return rather than return in response to the resolution of a civil war. Since we cannot expect the rebels to anticipate the natural disaster and the return of refugees, the effect of the spontaneous return of refugees on violence is because of large population movements rather than strategic agents.

## **Empirical Strategy**

In brief, in my analysis the dependent variable of interest is change in violence. The independent variable of interest is number of refugee returning. I instrument for refugee return with incidents of natural disasters in the asylum country. I make use of controls for various factors in the origin country that are likely to cause violence, as well as controlling for facets of the country pair that suggest endogenous violence and refugee flows. The unit of analysis is origin country (indexed by  $i$ ), asylum country (indexed by  $j$ ), year (indexed by  $t$ )

More particularly, since the claim is not that refugee return causes increases in violence, but large influxes of population cause violence, I incorporate the critical requirement large

populations by interacting the natural disasters with the number of refugees from the country of origin in the asylum country in the year before the disaster. This number can be thought of as the potential returnees. Furthermore, since I am only interested in natural disasters that will change the conditions sufficiently to push refugees out of the asylum country, I restrict the instrument to years in which a natural disasters rendered at least 7,500 persons homeless. I will refer to this as a major natural disaster.

There are a handful of reason that the major natural disasters in asylum countries could cause increases in violence in civil wars in countries of origin. Therefore, to ensure the exclusion restrictions holds, I condition on these. More particularly, as discussed earlier since natural disasters are not confined by national borders, it is possible that the same natural disaster occurs in the country of origin and the country of asylum. A natural disaster in a country with a civil war may cause an increase in hostilities, thus I control for major natural disasters in the country of origin also. Armed conflict in the asylum country may also cause refugees to return differentially in response to the natural disaster. Therefore I control for armed conflict in the asylum country as well. One might worry that a natural disaster might increase the number of people from the asylum country fleeing to the civil war country, causing more than just the returnees to arrive. I control for refugee flows in the opposite direction, as well. Distance between the country of origin and country of asylum likely attenuates the effect. Thus, I control for distance between the country pair.

Last, I control for a number of characteristics that the literature suggests are determinants of violence. This includes time varying characteristics like how relatively wealthy the country is, whether the country is a democracy, the size of the population. It also include time invariant characteristics like how mountainous the country is, the ethnic fractionalization and the religious fractionalization of the country's population.

Thus, the naive model of interest is:

$$\Delta Violence_{it} = \beta_0 + \beta_1 Returnees_{ijt} + Controls_{ijt}\beta_2 + \eta_{ijt} \quad (1)$$

Thus, the instrumented model of interest is:

$$\begin{aligned} Returnees_{ijt} = & \gamma_0 + \gamma_1 Major\ Natural\ Disaster_{jt} \times Refugees_{ij,t-1} \\ & + \gamma_2 Major\ Natural\ Disaster_{jt} + \gamma_3 Refugees_{ij,t-1} + \nu_{ijt} \end{aligned} \quad (2)$$

$$\Delta Violence_{it} = \beta_0 + \beta_1 \hat{Returnees}_{it} + Controls_{ijt} + \eta_{ijt} \quad (3)$$

## The Data

In order to study the relationship between refugee flows and civil conflict, I construct a cross country dataset with an origin-asylum-year unit of analysis from 1976 to 2005. I combine data from the United Nations High Commission for Refugees (UNHCR) on refugee flows with data from the Peace Research Institute Oslo UCDP/PRIO Armed Conflict Dataset. For the incidence of natural disasters, I make use of the Centre for Research on the Epidemiology of Disasters' (CRED) International Disaster Database (EM-DAT). The following section explains important features of the data used.

### UNHCR Refugee Data

The UNHCR counts refugees as those who are recognized under the 1951 *UN Convention Relating to the Status of Refugees*, its corresponding 1967 *Protocol*, and the *OAU Convention Governing the Specific Aspects of Refugee Problems in Africa*. Refugees, according to the 1951 Convention are people “who [are] unable or unwilling to return to their country of origin

owing to a well-founded fear of being persecuted for reasons of race, religion, nationality, membership of a particular social group, ...or political opinion” (United Nations 1951). UNHCR also sometimes counts some refugees recognized by individual countries of asylum or under national laws. To be clear, all of the persons counted by UNHCR are fleeing some kind of violence or persecution. Their definition works well for me because the vast majority of UNHCR refugees who are counted are those fleeing a country with some kind of armed conflict.

UNHCR’s *Statistical Online Population Database* provides counts of refugees in each country of asylum by each country of origin. UNHCR’s data is compiled from multiple sources and counts refugees primarily using registries.<sup>3</sup> When numbers of refugees are not available for a country-pair-year, I substitute numbers of UNHCR assisted refugees. Counts are available beginning in 1975 (UNHCR 2011). When matching it with other available data, my panel extends to 2005. This time frame is long enough to consider a variety of conflict across the world.

Because the return of refugees data are available for far fewer years and countries, I define returnees based on the change from the preceding year refugee population. The refugee return variable is constructed from the year over year difference of refugee counts by origin and country of asylum from UNHCR. If the number of refugees decreased over the previous year in a country of asylum from a particular country of origin and the total number of refugees from that country of origin decreased from the previous year, the refugee return variable is the change in the number of refugees for the origin, asylum pair. If the number of refugees in the asylum country from the country of origin remained the same, increased, or the total number of refugees from the country of origin remained the same or increased, the variable is

---

<sup>3</sup>UNHCR does not include data from most refugees in Palestinian camps as this is the domain of United Nations Relief and Works Agency for Palestinian Refugees in the Near East (UNRWA) and likely undercounts refugees not living in camps, so I exclude Palestine. For a more detailed account of definitions and limitations, see <http://www.unhcr.org/45c06c662.html>.

coded zero.<sup>4</sup> For years in which counts of voluntarily repatriated refugees are available (the primary source of returnees), my measure has a reasonably high correlation, 0.71.

## Conflict Data

With the data on refugees, I pair information on internal armed conflict from PRIO (Gleditsch et al. 2002; Harbom and Wallensteen 2010). UCDP/PRIO defines an armed conflict as “a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths” (Harbom and Eriksson 2009). The dataset also makes a distinction (by intensity) between an armed conflict and a war by battle deaths. Over 1,000 battle deaths per year is a war, and fewer battle deaths is a conflict (Harbom and Eriksson 2009).

Specifically, I count the number of internal (type 3 and type 4) armed conflicts per country, per year and by intensity. Then I construct a change in violence variable. The variable is coded as 1 if one of two conditions is met: (1) if the number of internal armed conflicts increased, or (2) if the number of internal armed conflicts considered wars (at least 1,000 battle deaths per year) increased. If the number of internal armed conflicts and those considered wars remained the same, then the change in violence variable is coded zero. If at least one of these elements decreased and all of other elements remained the same, the change in violence variable is coded -1.<sup>5</sup>

## Natural Disasters

EM-DAT counts disasters if one of four conditions is met: (1) at least 10 deaths were reported; (2) at least 100 persons were affected; or there was (3) a declaration of national

---

<sup>4</sup>I include the condition of an overall decrease in refugees from the preceding year, to prevent counting refugees who moved from one asylum country to another asylum country.

<sup>5</sup>While there is battle death count data for many of the conflicts in the dataset, it is not available for all of them. Therefore, I use the conflict intensity data to determine changes in violence. Thus the variable can be thought of as a simplification of the underlying continuous quantity percent change in battle deaths.

emergency or (4) a call for international assistance (EM-DAT 2011). Natural disasters include earthquakes, volcanic eruptions, dry mass movement (including landslides), storms, floods, wet mass movements (including avalanches and mudslides), extreme temperature (hot and cold), droughts and wildfires. They further count the number of persons left homeless by the event. The major natural disaster indicator variable is coded 1 if at least 7,500 people were left homeless because of at least one natural disaster in a given country for the given year, and zero otherwise.

## Controls

I draw on other scholars' work on civil war for control variables. I make use of Fearon and Laitin (2003) data for the percent mountainous terrain of the country. I use (Alesina et al. 2003) data for ethnic and religious fractionalization. I use Gleditsch (2002) for population and real per capita GDP.<sup>6</sup> I use the Mayer and Zignago (2010) for a measure of distance between country pairs. Last, the asylum country war indicator is constructed using PRIO type 2 (interstate), type 3 and type 4 (internal conflicts).<sup>7</sup> The origin country natural disaster indicator is constructed just like the one for the asylum country. The exchange of refugees in both directions is coded 1 if there is an UNHCR year over year increase in refugees from the country pair, in the opposite direction.

## Summary Statistics

Tables 1 provide descriptive statistics of the data used in the analysis.

[Insert Table 1 about here.]

---

<sup>6</sup>I use this data because it is compiled from multiple sources and interpolated, reducing the number of missing data. The shortcoming is the most recent beta version of the data is only available through 2004. Gleditsch (2002) for a complete discussion of the merits of the data.

<sup>7</sup>There were no Type 1 (Colonial Wars) for the years of interest.

Table 2 breaks the descriptive statistics out by change in the level of violence in the civil conflict of the origin country for the data used. As should be the case the average incidence of natural disasters does not statistically differ across changes in violence.

[Insert Table 2 about here.]

## Results

This section explores the impact of refugee return on civil conflict intensity in detail. I begin by presenting the results of ordinary least squares analysis on the relationship between returnees and violence. Like anticipated the relationship is negative, although not statistically significant in most specifications. Then I present the two stage least squares analysis. The effect of refugee return on violence is a 3%-8% increase in violence for every 100,000 returnees. The positive effect is statistically significant and robust to a range of specifications. I conclude the section with some reasons to be cautious about the analysis.

Table 3 displays the results of the Ordinary Least Squares (OLS) regressions from models (1) and (2) and a revised model (2) that includes a control for democracies based on the Polity IV score (Marshall and Jaggers 2010).<sup>8</sup>

[Insert Table 3 about here.]

These reduced form results are suggestive of the traditional understanding in the literature of the relationship between refugee return and violence. Namely, refugees return to their country of origin when they perceive it is safe to do so, and refugee return is correlated with decreases in violence. Specifically the coefficient on the return relationship to violence is negative in the simplest regression (1) and insignificant. On the whole, the impact of refugee

---

<sup>8</sup>Polity IV does not provide data for a host of countries that its makers have determined are in transition or occupied. Since wars are often associated with transitions and occupations, this is less than desirable. This indicator is coded as one when the Polity Score is positive and zero otherwise.

return is consistently small. The return of 100,000 people is only associated with a 1.4% decrease in violence.

While the controls add some precision to the regression, the estimates from model (2) still suggest a negative but not significant relationship. The return of 100,000 people is still associated with a 1.4% decrease in violence. Adding, the polity control, does not change the sign, although the estimate drops considerably and on the whole, there is little precision in the estimates. At first glance, this would lead us to the conclusion that refugees return is associated with a decrease in violence, but ambiguously so, since none of these estimates are distinct from zero.

One concern with these results might be concerns about omitted variable bias. That the negative, statistically insignificant result is being driven by the absence of particular controls. Table 4 explores this relationship in more detail by examining the model with various sets of fixed effects.

[Insert Table 4 about here.]

The first column displays the analysis using fixed effects for seven regions of the world as designated by the World Bank. While this decreases the point estimate slightly, the results remain statistically indistinguishable from zero. The second column displays results for the analysis with origin country fixed effects. In this case, the estimate is a bit larger in magnitude 2.6% decrease in violence for every 100,000 returnees. This estimate is statistically significant at the 90% level. Finally, the third column presents the results from analysis including year fixed effects. Like the region fixed effects, this has little impact on the quantity of interest.

The Two Stage Least Squares analysis displays importantly different results. Now instead of a negative relationship, the relationship is positive. In these models, violence increases 3% for every 100,000 returnees.

[Insert Table 5 about here.]

More particularly, Table 5 shows first that the incidence of natural disaster interacted with the potential return population is a valid instrument for refugee return. Second, this model reveals a positive relationship between increased violence and refugee return. The last two columns of Table 5 make use of controls. The controls increase our estimate of the impact of returning refugees. The impact of 100,000 returned refugees is a 3% increase in violence, and the effect is statistically significant.

The difference in results between the ordinary least squares regression and the two stage least squares are remarkable. Not only does the sign of the effect change when the endogenous relationship between conflict and refugees flow is accounted for, but the instrumented estimates are twice the magnitude of the ordinary least squares estimates. These two stage least squares estimates allows us to identify that returnees not only increase violence, but the impact of population flows is larger than the OLS regression suggests.

## **Some Checks on the Analysis**

In this subsection I deal with two major concerns with the analysis presented thus far. First, is the concern that the results reported are because of misspecification of the model. I address this by presenting Ordered Probit and Two Stage analysis with Ordered Probit models. The results are similar to those from the Ordinary Least Squares and Two Stage Least Squares models. The second concern is that the exclusion restriction is violated and the increases in civil war violence are not due to the returnee flows but some other channel that connects the natural disaster in the asylum country to the violence in the country of origin that I do not take into account. To address this issue, I present the results of a placebo test.

A first concern with my analysis is a concern about misspecification. The dependent variable, change in civil war intensity is coded categorically – a decrease, no change, and an increase

in violence. The reason for choosing this coding is that it allows me to use a larger set of conflicts from the Armed Conflict Database, necessary for doing the actual analysis. These changes in violence intensity are a simplification of a continuous variable, what they capture is akin to percent changes in battle deaths year over year. If the battle death data were available for a larger set of conflicts, it would be the optimal solution to misspecification concern.<sup>9</sup>

In the absence of that data, there are still good methodological reasons for using a linear model rather than the Ordered Probit. The first reason is rather superficial, since the quantity of interest is something more like percent change in battle deaths, predictions outside the bounds of probability are not a problem. The second reason is that since I am doing two-stage analysis, the assumptions required for consistency of the estimates is more straightforward in the linear case. Third and finally, since I do some fixed effects analysis, making a comparison between the estimate with and without fixed effects only makes sense if the fixed effects estimates are unbiased, which would not be the case using an Ordered Probit.<sup>10</sup>

Nevertheless, the Ordered Probit and two stage Ordered Probit analysis results are consistent with those of the linear analysis. Table 6 displays the coefficients on refugee return in 12 different models.

[Insert Table 6 about here.]

The first line is directly from Table 3. The same analysis is repeated using two stage least squares to estimate the coefficients, then ordered probit. Finally, the coefficients are estimated

---

<sup>9</sup>In the detail of the Battle Death data from PRIO, many of the estimates for battle deaths are simply the 25 and 1,000 battle death thresholds that the conflict intensity variable reflects, providing no real additional data on change in violence. The new geocoded events datasets (ACLED or PRIO-GRID) may be another good source of data for violence. However, in this case these datasets do not overlap temporally sufficiently with other data for the analysis.

<sup>10</sup>See Angrist and Pischke (2008) and Greene and Hensher (2010) for discussions of the merits of these choices.

using a linear first stage and an Order Probit second stage. Note that the first stage is the same as reported in Table 5 for all the instrumented models.

The results of the Ordered Probit analysis is substantively similar to that of the linear analysis. Without an instrument, the effect of refugee return on violence is negative, and not statically distinct from zero. The linear estimates are about half those of the Ordered Probit in magnitude. In the instrumented version, the results are positive, small and statistically significant regardless of the model. The estimates from the linear analysis are smaller in magnitude and more precises, where as the Order Probit estimates are larger and less precise. Qualitatively, regardless of the methodological choice, between 3% and 8% increased violence occurs for every 100,000 returnees. This is consistent with the literature that attributes some violence to large population shifts alone.

A second consideration with my analysis is that one might think that there is something unaccounted for about natural disasters that causes increases in civil war violence in a neighboring country even in the absence of refugee flows (in other words, that natural disasters do not serve as a valid instrument for the analysis). If this were the case, natural disasters could account for the change in violence, but have nothing to do with returnees inciting the violence. To test this concern I constructed a dataset of country-pair-years in which the two countries are contiguous (Mayer and Zignago 2010), the first country has had a civil conflict in the preceding year, and the contiguous country did not receive refugees. With this set of countries, I examine the relationship between the contiguous country having a natural disaster and the change in intensity of the civil conflict in the country of origin controlling for both war in the contiguous country and a natural disaster in the origin country. The results of this comparison are in Table 7.

[Insert Table 7 about here.]

Table 7 suggests that there is a negative, but not significant, relationship between contiguous countries natural disasters and civil conflict, when there are no refugee flows. This would suggest that my instrument, using natural disasters may be biased. However, it would be biased downward since the relationship between the contiguous country's natural disaster and the civil conflict intensity is negative and therefore makes my causal story more difficult to prove. Since I still find a positive and significant result despite this potential bias, it is not a problem for my results.

## A Few Problems with the Data

There are good reasons to have confidence in the findings explored here. The results are robust to the consideration of a number of theoretically relevant controls and to a set of reasonable model specifications. Furthermore, the placebo test reinforces the value of the instrument in this case. The instrument accomplishes two things, it disaggregates violence attributable to moving violent individuals, like rebels, and violence attributable to the large population shift. Furthermore the magnitude of the finding is consistent with existing theory, in the sense that violence is associated not with small population movements, but very large population movements that can constitute a shock. In many places in the world a population change of 100,000 people in a year is sufficiently large that we can think of it as shock. This is the order of magnitude of change in population required to see noticeable effects.

However, like much of the cross-country analysis in the area of conflict, this project is subject to a number of data challenges.<sup>11</sup> The results presented here should be interpreted with consideration for the critiques leveled against comparable uses of the data before. These kinds

---

<sup>11</sup>Measurement is noisy and there is a lot of missing data. There is every reason to think that these problems are not spurious but directly related to the quantities of interest in the analysis. Specifically, it is harder to count people, when they arrive more quickly, like in refugee crisis. It is difficult to know how much violence and death takes place in civil wars, especially those where there is less monitoring at the earlier part of the data set. Finally, missing data on population, the strength of institutions, and macro economic measure like GDP is directly related to the civil wars and population movements being studied.

of problems notwithstanding, there are a few exclusions from the cross-country dataset that are qualitatively important. First, information about Palestinian refugees are all together excluded from the data. The UNHCR numbers on Palestinian refugees are incomplete because they do not include any of the refugees under the protection of the UNRWA. Second, UNHCR counts refugees from Western Sahara and from Tibet separately from any recognized country. Since these regions are not recognized generally in the international community as states in their own right, their unit of analysis differs from the rest of the dataset. Furthermore, there is not sufficient matching data to include them in my analysis. I can think of no reason to believe that the causal relationship that I have explored above would be different for these Palestine, Western Sahara or Tibet. However, while it may be the case that the relationship between refugee return and increased violence is the same for Palestinian, Tibetan or Sahrawi refugees, these represent some of the most long standing refugee crises outside of sub-Saharan Africa. A study with microdata of any one of these groups of refugees, their movements and violence, would be a good compliment to the study presented here.

## **Discussion**

My exploration of the empirical relationship between refugee return and increased violence in civil conflict reveals that large refugee return causes increases in violence in armed civil conflict. Violence increases about 3-8% for every 100,000 returnees. The measurable increase in violence in the context of civil war gives credence to concerns over large population movements destabilizing their destination, especially when the stability of the destination is fragile prior to the arrival of migrants. Furthermore, this finding is distinct from the literature that connects spreading violence with militant refugees and rebels hidden amongst refugees in exile. This finding is more consistent with sociological literature that finds violence increases even when there is no civil war and there are no rebels or agents strategically employing organized violence. Rather it is more consistent with large population shifts straining the

existing structures that are maintaining order.

From the perspective of political economy and identifying causal mechanisms, this explanation that existing structures for maintaining order are strained is far too general. In the literature on civil war, state failure and population shifts and destabilization often have two varieties. The first hinges on ethnic or religious differences (Fearon 1994; Lake and Rothchild 1996; Sambanis 2001) and the second is an institutional story where regardless of the underlying reason, the state fails at providing security or alternatives to rebellion when strained (Humphreys 2005; Miguel et al. 2004). In an attempt to think more about these two categories of potential mechanisms, I look for heterogeneity in the effects on violence along two dimensions: the presence of an ethnic or religious war (Sambanis 2001) in the country of origin and the strength of institutions (Marshall and Jaggers 2010) in the country of origin. Both of these empirical inquiries yield largely null findings.<sup>12</sup>

I do a series of analyses using fixed effects in the two stage least squares models. The results in these models reveal heterogeneity in the effects of refugee return on violence. This set of empirical findings mirrors the state of the theoretical literature, in the sense that there are country specific characteristics that are unaccounted for that explain why large refugee return sometimes prompts increased violence and sometimes does not. Additional theoretical development is necessary before more empirical work can be done. To the extent the results point to a probable explanation, it likely lies in an institutional kind of explanation rather than an ethnic differences explanation. Such an explanation needs to be able to connect large (and not small) refugee return with an institutional problem that leads to increased rebellion.

In the absence of a testable causal mechanism, this article goes far in establishing the need

---

<sup>12</sup>It's possible that this is due to reduced sample size in the case of analyzing ethnic wars. However, it may also be that large population shifts and ethnic difference explanations do not account for the increase in violence identified here. In the case of strength of institutions, the empirical problem is almost certainly missing data for occupied and transitional states. These states are among those that are experiencing increased violence, and likely driving the results.

for one. It is not just the case that militant refugees or rebels among refugees can cause the spread of violence. It is also the case that in the context of fragile stability or an ongoing civil war, that large population movements, like those in refugee crises, themselves can cause violence. This project lends credibility to the claims of asylum countries, like Lebanon that large refugee populations threaten to destabilize already fragile political situations, and can lead to increased violence.

## References

- Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003). Fractionalization. *Journal of Economic Growth* 8(2), 155–194.
- Angrist, J. and J. Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Baum, C. F., M. E. Schaffer, and S. Stillman (2002). IVREG2: Stata module for extended instrumental variables/2SLS and GMM estimation. Statistical Software Components, Boston College Department of Economics.
- EM-DAT (2011). EM-DAT: The OFDA/CRED International Disaster Database. <http://www.emdat.be>. Université Catholique de Louvain, Brussels (Belgium).
- Fearon, J. and D. Laitin (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review* 97(1).
- Fearon, J. D. (1994). Ethnic war as a Commitment Problem. *Mimeo*.
- Gleditsch, K., I. Salehyan, and K. Schultz (2008, April). Fighting at home, fighting abroad: How civil wars lead to international disputes. *Journal of Conflict Resolution* 52(4), 479–506.
- Gleditsch, K. S. (2002). Expanded Trade and GDP Data. *Journal Of Conflict Resolution* 46(5), 712–724.
- Gleditsch, N. P., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand (2002). Armed Conflict 1946-2001: A New Dataset. *Journal of Peace Research* 39(5), 615–637.
- Greene, W. and D. Hensher (2010). *Modeling Ordered Choices: A Primer*. Cambridge University Press.

- Harbom, L. and M. Eriksson (2009). *UCDP / PRIO Armed Conflict Dataset Codebook v4.2009*.
- Harbom, L. and P. Wallensteen (2010, July). Armed Conflicts, 1946–2009. *Journal of Peace Research* 47(4), 501–509.
- Humphreys, M. (2005). Natural resources, conflict, and conflict resolution uncovering the mechanisms. *Journal of conflict resolution* 49(4), 508–537.
- ICG (1999). Towards Destabilization? The Kosovo crisis takes its toll on Macedonia. *ICG Report 67*. <http://www.crisisgroup.org/~media/Files/europe/Macedonia%206.pdf>, 21 May 1999.
- ICG (2013). Too Close For Comfort: Syrians in Lebanon. *Middle East Report 141*. <http://www.crisisgroup.org/~media/Files/Middle%20East%20North%20Africa/Iraq%20Syria%20Lebanon/Lebanon/141-too-close-for-comfort-syrians-in-lebanon.pdf>, 13 May 2013.
- Lake, D. A. and D. Rothchild (1996). Containing Fear: The Origins and Management of Ethnic Conflict. *International Security* 21(2), 41–75.
- Lischer, S. K. (2006). *Dangerous sanctuaries: refugee camps, civil war, and the dilemmas of humanitarian aid*. Cornell University Press.
- Lucey, J. F. (1914). *Lucey to Hoover, describing confused situation in Belgium as reported by local officials, refugees, and others*. Stanford University Press.
- Marshall, M. G. and K. Jaggers (2010). Polity IV Dataset (Computer file; version p4v2010).
- Mayer, T. and S. Zignago (2010). Notes on CEPII’s distances measures: The GeoDist database.

- Miguel, E., S. Satyanath, and E. Sergenti (2004, August). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *The Journal of Political Economy* 112(4), pp. 725–753.
- Mikati, N. (2014). Lebanon PM: My country cannot cope with the Syrian refugee crisis. *The Daily Telegraph*. <http://www.telegraph.co.uk/news/worldnews/middleeast/lebanon/10587174/Lebanon-PM-My-country-cannot-cope-with-the-Syrian-refugee-crisis.html>, 21 Jan 2014.
- Roodman, D. (2011). Fitting Fully Observed Recursive Mixed-Process Models with `cmp`. *Stata Journal* 11(2), 159–206.
- Salehyan, I. (2008). The Externalities of Civil Strife: Refugees as a Source of International Conflict. *American Journal of Political Science* 52(4), 787–801.
- Salehyan, I. and K. Gleditsch (2006). Refugees and the spread of civil war. *International Organization* 60(02), 335–366.
- Saleyan, I. (2011). *Rebels without Borders: Transnational Insurgencies in World Politics*. Cornell University Press.
- Sambanis, N. (2001). Do Ethnic and Nonethnic Civil Wars have the Same Causes? A Theoretical and Empirical Inquiry (Part 1). *Journal of Conflict Resolution* 45(3), 259–282.
- Sampson, R. J. and J. H. Laub (2005). A life-course view of the development of crime. *The Annals of the American Academy of Political and Social Science* 602(1), 12–45.
- Shaw, C. R. and H. D. McKay (1942). Juvenile delinquency and urban areas.

UNHCR (2011). Statistical Online Population Database. United Nations High Commissioner for Refugees (UNHCR). <http://www.unhcr.org/statistics/populationdatabase> [Date extracted: 18 April 2011].

United Nations (1951). The 1951 Convention Relating to the Status of Refugees.

Yacoubian, M. (2014). Renewed Conflict in Lebanon. *Contingency Planning Memorandum 22*.

Table 1: Summary Statistics Full Sample

	Full Sample		
	Mean	SD	N
Refugee Return (1,000)	0.863	18.197	30,781
<u>Origin Country</u>			
Natural Disaster	0.147	0.354	30,781
Log % Mountainous Terrain	2.352	1.370	30,781
Ethnic Fractionalization	0.543	0.251	30,781
Religious Fractionalization	0.437	0.229	30,781
Polity Score (-10, 10)	0.203	6.101	29,439
Population, Lagged, (1,000)	49.9	160.9	30,781
Real GDP per Capita, Lagged (1,000)	3.980	4.848	30,781
<u>Asylum Country</u>			
War Occurring	0.131	0.338	30,781
Natural Disaster	0.089	0.285	30,781
<u>Country Pair</u>			
Refugees in Both Directions	0.066	0.248	30,781
Distance between Countries (km)	5,293	3,974	30,781

Table 2: Summary Statistics by Violence

	Decrease		No Change		Increase	
	Mean	SD	Mean	SD	Mean	SD
Refugee Return (1,000)	1.616	23.280	0.690	16.000	1.299	25.368
<u>Origin Country</u>						
Natural Disaster	0.186	0.389	0.145	0.353	0.116	0.320
Log % Mountainous Terrain	2.600	1.313	2.254	1.377	2.793	1.259
Ethnic Fractionalization	0.614	0.231	0.521	0.253	0.628	0.218
Religious Fractionalization	0.446	0.218	0.435	0.234	0.444	0.204
Polity Score (-10, 10)	0.366	4.917	0.474	6.354	1.197	5.016
Population, Lagged, (1,000)	55.4	152.2	47.8	162.9	58.7	154.2
Real GDP per Capita, Lagged (1,000)	2.478	3.257	4.435	5.140	2.325	3.131
<u>Asylum Country</u>						
War Occurring	0.150	0.357	0.124	0.330	0.162	0.369
Natural Disaster	0.098	0.297	0.087	0.281	0.095	0.294
<u>Country Pair</u>						
Refugees in Both Directions	0.084	0.277	0.061	0.240	0.079	0.270
Distance between Countries (km)	5,124	3,706	5,346	4,053	5,089	3,653
Observations*	3,548		3,548		3,548	

\*Except for Polity for which  $N$  is 3,417 , 22,978, and 3,044 for each category respectively.

Table 3: Ordinary Least Squares (DV: Change in Violence)

	Model (1) Coeff ( $\hat{\beta}$ ) <i>SE</i>	Model (2) Coeff ( $\hat{\beta}$ ) <i>SE</i>	Model (2') Coeff ( $\hat{\beta}$ ) SE
Refugee Return (1,000)	-0.00014 <i>0.00013</i>	-0.00014 <i>0.00012</i>	-0.00014 <i>0.00012</i>
Asylum Country War		0.01080 <i>0.00737</i>	0.01053 <i>0.00717</i>
Flows in Both Directions		-0.01494 <i>0.01185</i>	-0.01516 <i>0.01195</i>
Significant Natural Disaster (origin)		-0.07593 <i>0.05218</i>	-0.07601 <i>0.05221</i>
Religious Fractionalization (origin)		-0.00721 <i>0.03909</i>	-0.00648 <i>0.03925</i>
Ethnic Fractionalization (origin)		0.03278 <i>0.03561</i>	0.03138 <i>0.03576</i>
Log % Mountainous Terrain		0.01222 * <i>0.00700</i>	0.01227 * <i>0.00699</i>
Population, Lagged (1,000)		0.00006 <i>0.00004</i>	0.00006 <i>0.00004</i>
Real GDP per Capita (1,000)		0.00038 <i>0.00092</i>	0.00050 <i>0.00098</i>
Democracy Indicator (origin)			-0.00572 <i>0.02475</i>
Distance Between Countries		-7.41E-08 <i>6.35E-07</i>	-2.55E-08 <i>6.97E-07</i>
Constant	0.00338 <i>0.00750</i>	-0.04353 * <i>0.02290</i>	-0.04065 0.02857
Number of Observations	44,712	30,781	30,781
Clusters	200	153	153

Notes: Standard errors are clustered at the origin country level; \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$

Table 4: Fixed Effects Analysis (DV: Change in Violence)

	Coeff <i>SE</i>	Coeff <i>SE</i>	Coeff <i>SE</i>
Refugee Return (1,000)	-0.00011 <i>0.00009</i>	-0.00026* <i>0.00015</i>	-0.00010 <i>0.00009</i>
Controls	Y	Y	Y
Region Fixed Effects	Y	N	N
Origin Country Fixed Effects	N	Y	N
Year Fixed Effects	N	N	Y
Number of Observations	30,781	30,781	30,781
Clusters	153	153	153

Notes: Standard errors are clustered at the origin country level; \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$

Table 5: 2SLS (IV: Natural Disaster, DV: Change in Violence)

	1st Stage Coeff ( $\hat{\gamma}$ ) <i>SE</i>	2nd Stage Coeff ( $\hat{\beta}$ ) <i>SE</i>	1st Stage Coeff ( $\hat{\gamma}$ ) <i>SE</i>	2nd Stage Coeff ( $\hat{\beta}$ ) <i>SE</i>	
Refugee Return (1,000)		0.00030 <i>0.00028</i>		0.00030 <i>0.00010</i>	***
Nat Disaster X Potencial Returnees	0.0534 * <i>0.0319</i>		0.0560 * <i>0.0319</i>		
Natural Disaster (Asylum)	-0.7234 ** <i>0.3420</i>		-0.8712 ** <i>0.4255</i>		
Potential Returnees (1,000)	0.0781 *** <i>0.0278</i>		0.0760 *** <i>0.0283</i>		
Flows in Both Directions			2.0072 <i>1.3098</i>	-0.0168 <i>0.0120</i>	
Significant Natural Disaster (origin)			0.0579 <i>0.2866</i>	-0.0760 <i>0.0521</i>	
Religious Fractionalization (origin)			0.5482 * <i>0.3276</i>	-0.0072 <i>0.0388</i>	
Ethnic Fractionalization (origin)			-0.4010 <i>0.4599</i>	0.0326 <i>0.0353</i>	
Population, Lagged (1,000)			-0.0004 <i>0.0003</i>	0.0001 <i>0.0000</i>	*
Real GDP per Capita (1,000)			-0.0272 <i>0.0207</i>	0.0004 <i>0.0009</i>	
Asylum Country War			-0.0437 <i>0.2574</i>	0.0100 <i>0.0072</i>	
Log % Mountainous Terrain			-0.0454 <i>0.0427</i>	0.0121 <i>0.0069</i>	*
Distance Between Countries			3.56E-06 <i>1.37E-05</i>	-1.34E-08 <i>6.33E-07</i>	
Constant	0.1289 <i>0.1038</i>	0.0031 <i>0.0021</i>	0.2361 <i>0.2932</i>	-0.0436 <i>0.0227</i>	*
F-Stat	333.38		376.68		
Number of Observations	44712		30781		
Clusters	153		153		

Notes: Standard errors are clustered at the origin country level; \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ , 2SLS is calculated using the ivreg2 module (Baum et al. 2002)

Table 6: Comparison of Coefficients on Refugee Return by Model (DV: Change in Violence)

<u>OLS</u>					
Coefficient	-0.00014	-0.00014		-0.00014	
Standard Errors	0.00013	0.00012		0.00012	
<u>Linear Second Stage</u>					
Coefficient	0.00030	0.00030	***	0.00030	***
Standard Errors	0.00028	0.00010		0.00010	
<u>Ordered Probit</u>					
Coefficient	-0.00037	-0.00034		-0.00034	
Standard Errors	0.00035	0.00031		0.00031	
<u>Ordered Probit Second Stage</u>					
Coefficient	0.00090	0.00083	**	0.00081	**
Standard Errors	0.00062	0.00035		0.00037	
Controls	N	Y		Y	
Polity Democracy Indicator	N	N		Y	
Observations	44,712	30,781		30,781	
Clusters	199	153		153	

Notes: Standard errors are clustered at the origin country level; \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ , 2SLS is calculated using the ivreg2 module (Baum et al. 2002), Instrumented Ordered Probit calculated using cmp module (Roodman 2011)

Table 7: Contiguous Countries, Impact of Natural Disasters

	DV: Change In Violence		
	Coefficient	SE	
Natural Disaster (Contiguous Country)	-0.0600	0.0559	
Natural Disaster (Origin Country)	-0.0462	0.0572	
War (Contiguous Country)	0.0484	0.0328	
Constant	0.1959	0.0387	***
Number of Observations	2112		
Clusters	80		

Notes: Standard errors are clustered at the origin country level; \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$